(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 October 2000 (26.10.2000)

PCT

(10) International Publication Number
WO 00/63435 A2

(51) International Patent Classification[7]: C12Q 1/68, C12N 15/70, 5/10, C07K 14/47, 16/18

(21) International Application Number: PCT/US00/10076

(22) International Filing Date: 14 April 2000 (14.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/129,763	15 April 1999 (15.04.1999)	US
60/156,924	28 September 1999 (28.09.1999)	US
09/548,589	13 April 2000 (13.04.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/129,763 (CIP)
Filed on	15 April 1999 (15.04.1999)
US	60/156,924 (CIP)
Filed on	28 September 1999 (28.09.1999)
US	09/548,589 (CIP)
Filed on	13 April 2000 (13.04.2000)

(71) Applicant *(for all designated States except US)*: **CURAGEN CORPORATION** [US/US]; 11th Floor, 555 Long Wharf Drive, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GOULD-ROTHBERG, Bonnie, E.** [CA/US]; 1701 Moose Hill Road, Guilford, CT 06437 (US). **DIPIPPO, Vincent, A.** [US/US]; 156 North Atwater Street, East Haven, CT 06512 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovksy, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

(48) Date of publication of this corrected version:
10 May 2001

(15) Information about Correction:
see PCT Gazette No. 19/2001 of 10 May 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/63435 A2

(54) Title: METHOD OF IDENTIFYING TOXIC AGENTS USING DIFFERENTIAL GENE EXPRESSION

(57) Abstract: Disclosed are methods of identifying toxic agents, *e.g.*, hepatotoxic agents, using differential gene expression. Also disclosed are novel nucleic acid sequences whose expression is differentially regulated by troglitazone.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12Q 1/68, C12N 15/70, 5/10, C07K 14/47, 16/18 — A2

(11) International Publication Number: WO 00/63435

(43) International Publication Date: 26 October 2000 (26.10.00)

(21) International Application Number: PCT/US00/10076

(22) International Filing Date: 14 April 2000 (14.04.00)

(30) Priority Data:

60/129,763	15 April 1999 (15.04.99)	US
60/156,924	28 September 1999 (28.09.99)	US
Not furnished	13 April 2000 (13.04.00)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/129,763 (CIP)
Filed on	15 April 1999 (15.04.99)
US	60/156,924 (CIP)
Filed on	28 September 1999 (28.09.99)
US	Not furnished (CIP)
Filed on	13 April 2000 (13.04.00)

(71) Applicant *(for all designated States except US)*: CURAGEN CORPORATION [US/US]; 11th Floor, 555 Long Wharf Drive, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: GOULD-ROTHBERG, Bonnie, E. [CA/US]; 1701 Moose Hill Road, Guilford, CT 06437 (US). DIPIPPO, Vincent, A. [US/US]; 156 North Atwater Street, East Haven, CT 06512 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovksy, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: METHOD OF IDENTIFYING TOXIC AGENTS USING DIFFERENTIAL GENE EXPRESSION

(57) Abstract

Disclosed are methods of identifying toxic agents, *e.g.*, hepatotoxic agents, using differential gene expression. Also disclosed are novel nucleic acid sequences whose expression is differentially regulated by troglitazone.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHOD OF IDENTIFYING TOXIC AGENTS USING DIFFERENTIAL GENE EXPRESSION

FIELD OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides and in particular to the identification of hepatotoxic agents in liver tissue using differential gene expression.

BACKGROUND OF THE INVENTION

An unfortunate drawback associated with otherwise promising drugs is that they induce unwanted side effects as well as their intended therapeutic effects. Often, these side effects do not become apparent until the drug has entered, or even completed, clinical trials. For example, troglitazone has recently been used to treat noninsulin dependent type II diabetes. In spite of its demonstrated effectiveness in alleviating symptoms associated with diabetes, significant side effects have been associated with this agent. In particular, it has been reported to result in liver damage in a subset of patients to which it is administered.

Hepatotoxicity associated with administration of troglitazone can range in severity from mild idiosyncratic reactions to severe hepatocellular necrosis, subsequent liver failure and death. Clinical manifestations of troglitazone-associated toxicity can include, *e.g.*, jaundice, nausea, abdominal pain, low-grade fever, malaise, vomiting, hematuria, lower-extremity edema and fatigue. Histopathological changes include, *e.g.*, cholestasis, hepatocellular necrosis, fibrosis, bile duct proliferation, and acute and chronic inflammation. Other changes associated with troglitazone-associated toxicity can include, *e.g.*, significant increases in bilirubin, alkaline phosphatase, alanine aminotransferase, and aspartate aminotransferase levels, and increased prothrombin time.

SUMMARY OF THE INVENTION

The invention is based in part on the discovery that certain nucleic acids are differentially expressed in liver tissue of animals treated with troglitazone. These differentially expressed nucleic acids include novel sequences and nucleic acids sequences that, while previously described, have not heretofore been identified as troglitazone responsive.

In various aspects, the invention includes methods of method of screening a test agent for toxicity, *e.g.*, hepatotoxicity. For example, in one aspect, the invention provides a method of identifying a hepatotoxic agent by providing a test cell population comprising a cell capable of expressing one or more nucleic acids sequences responsive to troglitazone, contacting the test cell population with the test agent and comparing the expression of the nucleic acids sequences in the test cell population to the expression of the nucleic acids sequences in a reference cell population. An alteration in expression of the nucleic acids sequences in the test cell population compared to the expression of the gene in the reference cell population indicates that the test agent is hepatotoxic.

In a further aspect, the invention provides a method of assessing the hepatotoxicity of a test agent in a subject. The method includes providing from the subject a cell population comprising a cell capable of expressing one or more troglitazone-responsive genes, and comparing the expression of the nucleic acids sequences to the expression of the nucleic acids sequences in a reference cell population that includes cells from a subject whoseexposure status to a hepatotoxic agent is known. An alteration in expression of the in the test cell population compared to the expression of the nucleic acids sequences in the reference cell population indicates the hepatotoxicity of the test agent in the subject.

Also provided are novel nucleic acids, as well as their encoded polypeptides, whose expression is responsive to the effects of troglitazone.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based in part on the discovery of changes in expression patterns of multiple nucleic acid sequences in rodent liver cells following exposure to troglitazone, a hepatotoxic compound. This compound is described in Spencer et al., Drugs 54(1): 89-102 (1997), whose contents are incorporated herein in their entirety.

The differentially expressed nucleic acids were identified by administering troglitazone to male 12 week old Sprague Dawley or Wistar Kyoto rats 617 mg/kg/day p.o. in QD dosing for 3 days. Control animals received dimethyl sulfoxide. The animals were sacrificed 24 hours following the last dose. Liver tissue was dissected from the animals, and total RNA was recovered from the dissected tissue. cDNA was prepared and the resulting samples were processed through using GENECALLING™ differential expression analysis as described in U. S. Patent No. 5,871,697 and in Shimkets et al., Nature Biotechnology 17:798-803 (1999). The contents of these patents and publications are incorporated herein by reference in their entirety.

819 gene fragments were initially found to be differentially expressed in rat liver tissue in response troglitazone in both Sprague Dawley and Wistar Kyoto rats. Genes fragments whose expression levels were modulated greater than $\pm$ 3-fold were selected for further analysis.

A summary of the sequences analyzed are presented in Table 1. Differential expression of 169 of the gene fragments was confirmed using a unlabeled oligonucleotide competition assay as described in Shimkets et al., Nature Biotechnology 17:198-803. The 169 single nucleic acid sequences identified herein, are referred to herein as HEPATOX 1-169. Column 6 of Table 1, entitled "Function", lists the type of classification assigned for the protein, based on its function.

Twenty-one sequences (HEPATOX: 1-21) represent novel rat genes for which the sequence identity to sequences found in public databases is either high (i.e., > 90%, observed for 5 fragments), moderate (i.e., between about 70% and about 90%, observed for 15 genes), or low (i.e., less than 70%, observed for 2 genes suggesting a putative homology.

The 148 other sequenced identified have been previously described. For some of the novel sequences (i.e., HEPATOX: 1-21), a cloned sequence is provided along with one or more additional sequence fragments (*e.g.*, ESTs or contigs) which contain sequences

substantially identical to, the cloned sequence. Also provided is a consensus sequences which includes a composite sequence assembled from the cloned and additional fragments. For a given HEPATOX sequence, its expression can be measured using any of the associated nucleic acid sequences may be used in the methods described herein. For previously described sequences (HEPATOX:22-169) database accession numbers are provided. This information allows for one of ordinary skill in the art to deduce information necessary for detecting and measuring expression of the HEPATOX nucleic acid sequences.

The troglitazone -responsive nucleic acids discussed herein include the following:

TABLE 1

Description of Sequence	GenBank Accession No.	PPARL Effect on Transcript Level		HEPATOX Assignment	SEQ ID NO	Function
		Sprague Dawley	Wistar-Kyoto			
Novel Genes						
Novel gene fragment, 185 bp, 84% SI to mouse EST AA255082	N/A	+4	+4	1	1	9.: UNKNOWN FUNCTION
Novel gene fragment, 56 bp, 84% SI to human clone HEP05890 [AK000303]	N/A	+2	+2	2	2	9.: UNKNOWN FUNCTION
Novel gene fragment, 572 bp, 71% SI to human clone CIT-HSPC_417O20 [AC008485]	N/A	-3	-4	3	3 & 4	9.: UNKNOWN FUNCTION
Novel gene fragment, 1770 bp, 65% SI to human BAC956 [U85199]	N/A	-3	-3	4	5 & 6	9.: UNKNOWN FUNCTION
Novel gene fragment, 755 bp, 86% SI to human clone DKFZp434B2411[AL137597]	N/A	-3	-7	5	7 & 8	9.: UNKNOWN FUNCTION
Novel gene fragment, 675 bp, 80% SI to a human bone marrow secreted protein designated BMS46 in patent database [Z36229]	N/A	-5	±1.0	6	9 & 10	4.: BASIC METABOLISM
Novel gene fragments, 874 bp, 89% SI to mouse bacteria binding macrophage receptor MARCO [U18424]	N/A	OFF	-4	7	11, 12, 13, 14 &15	9.: UNKNOWN FUNCTION
Novel gene fragment, 546 bp, 82% SI to mouse EST AW107981 (which is similar to human bile acid CoA: amino acid N-acyltransferase (O08833))	N/A	-3	-3	8	16 & 17	9.: UNKNOWN FUNCTION
Novel gene fragment, 407 bp, 68% SI to human bile acid CoA: amino acid N-acyltransferase [L34081]	N/A	0	-2	9	18	4.: BASIC METABOLISM
Novel gene fragment, 1149 bp, 82% SI to human docking protein (signal recognition particle receptor) [X06272]	N/A	-2	-3	10	19 & 20	9.: UNKNOWN FUNCTION
Novel gene fragment, 961 bp, 73% SI to human androgen induced protein (AIG-1) [AF153605]	N/A	+3	+4	11	21 & 22	9.: UNKNOWN FUNCTION
Novel gene fragment, 667 bp, 99% SI to mouse Sui1 [Z50159]	N/A	-3	-3	12	23 & 24	1.2.5.: TRANSLATION FACTORS
Novel gene fragment, 580 bp, 93% SI to mouse dihydrolipoamide dehydrogenase (Dld) [U73445]	N/A	+4	+2	13	25 & 26	4.5.4.: PYRUVATE FAMILY
Novel gene fragment, 327 bp, 77% SI to human apolipoprotein B-100 [M15421]	N/A	+3	+4	14	27 & 28	4.11.1.: EXTRACELLULAR TRANSPORT
Novel gene fragment, 650 bp, 79% SI to human putative progesterone binding protein [AJ002030]	N/A	+4	+5	15	29 & 30	4.11.2.: PLASMA MEMBRANE SHUTTLING
Novel gene fragment, 499 bp, 90% SI to mouse interferon regulatory factor 7 (mirf7) [U73037]	N/A	±1.0	-6	16	31 & 32	6.: EXTRACELLULAR ENVIRONMENTAL REGULATION
Novel gene fragment, 315 bp, 73% SI to human factor XI (blood coagulation factor) [M13142]	N/A	±1.0	-2	17	33	6.: EXTRACELLULAR ENVIRONMENTAL REGULATION
Novel gene fragment, 760 bp, 81% SI to mouse glycerol-3-phosphate acyltransferase [M77003]	N/A	±1.0	+2	18	34 & 35	4.: BASIC METABOLISM
Novel gene fragment, 498 bp, 88% S.I. to human delta-5 desaturase [AF199596]	N/A	+4	+3	19	36 & 37	9.: UNKNOWN FUNCTION
Novel gene fragment, 1083 bp, 96% SI to rat mitochondrial genome [X14848]	N/A	+4	+8	20	38 & 39	9.: UNKNOWN FUNCTION
Novel gene fragment, 376 bp, 81% SI to mouse liver alpha-amylase [V00719]	N/A	±1.0	-2	21	40 & 41	4.: BASIC METABOLISM

TABLE 1

Molecular Toxicology Marker Genes Set I (Monooxygenases)						
Cytochrome p450e phenobarbital inducible (CYP2B2)	M37134/ J00720/ AH002162	+6	NEW	22	N/A	4.9.1.: DETOXIFICATION: MONOOXYGENASES
Cytochrome p450e-L (CYP2B2)	M34452	+3	NEW	23	N/A	4.9.1.: DETOXIFICATION: MONOOXYGENASES
Cytochrome p450 (1) variant	K01721	+4	+45	24	N/A	4.9.1.: DETOXIFICATION: MONOOXYGENASES
PPAR Ligand Activated Genes						
Malic enzyme	M26581/ M26585	+4	+4	25	N/A	4.1.1.: FATTY ACID SYNTHESIS
Medium-chain acyl-CoA dehydrogenase (ACADM, MCAD, MCADH)	J02791	+4	+4	26	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION
Long-chain acyl CoA synthetase (palmitoyl-CoA ligase) (fatty acid coenzyme A ligase, long-chain (FACL1))	D90109	+4	+23	27	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION
Acyl-CoA oxidase	J02752	+2	+5	28	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
Peroxisomal enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase	K03249	+5	NEW	29	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
CYP 4A1 [Cytochrome p452]	X07259	+6	+5	30	N/A	4.1.2.3.: MICROSOMAL OMEGA OXIDATION
CYP 4A2	M57719	+2	+5	31	N/A	4.1.2.3.: MICROSOMAL OMEGA OXIDATION
CYP 4A3	M33936	+2	+6	32	N/A	4.1.2.3.: MICROSOMAL OMEGA OXIDATION
(Mitochondrial) 3-hydroxy-3-methylglutaryl CoA synthase	M33648	+3	+1.5	33	N/A	4.1.3.: KETONE BODY METABOLISM
Phosphoenolpyruvate carboxykinase	K03248	-2	-3	34	N/A	4.3.1.: GLYCOLYSIS/GLUCONEOGENESIS
Acute Phase Protein Genes/ Immune Response Genes						
Beta-galactoside alpha 2,6-sialyltransferase	M18769	+2	+3	35	N/A	1.4.: POSTTRANSLATIONAL MODIFICATION
Cathepsin B	X82396	-2	-3	36	N/A	1.5.1.: PROTEOLYSIS
Cathepsin C	D90404	+2	+5	37	N/A	1.5.1.: PROTEOLYSIS
Cathepsin S	L03201	-2	-2	38	N/A	1.5.1.: PROTEOLYSIS
Plasma proteinase inhibitor alpha-1-inhibitor III	J03552	+3	+2	39	N/A	1.5.1.1.: PROTEOLYSIS INHIBITOR
Alpha-1-protease inhibitor	D00675	-2	-3	40	N/A	1.5.1.1.: PROTEOLYSIS INHIBITOR
Hemopexin (beta-glycoprotein)	M62642	-2	-8	41	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Ceruloplasmin	L33869	-1.5	-2	42	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Haptoglobin (Hp)	M34230	+3	+2	43	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Serum amyloid P (SAP)	X55761	-2	-2	44	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Complement factor I (CFI)	Y18965	-2	-3	45	N/A	6.2.1.1.: COMPLEMENT COMPONENTS
Complement protein C1q beta chain	X71127	-2	-2	46	N/A	6.2.1.1.: COMPLEMENT COMPONENTS
Furin	X55660	-2	-2	47	N/A	6.2.1.1.: COMPLEMENT COMPONENTS
Pre-pro-complement C3	X52477	-2.5	-2	48	N/A	6.2.1.1.: COMPLEMENT COMPONENTS
Complement C8 beta (C8b)	U20194	-2	-2	49	N/A	6.2.1.1.: COMPLEMENT COMPONENTS
MHC-associated invariant chain gamma	X13044	-2	-3	50	N/A	6.2.2.1.: MHC PROTEINS
C-reactive protein	M83176	+2	NEW	51	N/A	6.2.4.: T-CELL PROCESSING AND PRESENTATION
Interleukin 6 receptor ligand binding chain	M58587	-3	-4	52	N/A	6.2.7.: MACROPHAGE/ PMN MEDIATED CYTOTOXICITY
Tumor necrosis factor receptor (TNF receptor)	M63122	-3	-3	53	N/A	6.2.7.: MACROPHAGE/ PMN MEDIATED CYTOTOXICITY

TABLE 1

Stress Response Genes						
Heat shock protein 60	X54793	+2	+6	54	N/A	1.3.1.: MOLECULAR CHAPERONE
Lysyl hydroxylase	L25331	-4	-4	55	N/A	1.4.: POSTTRANSLATIONAL MODIFICATION
Metallothionein-1 (and 2)	J00750/ M11794_1	-6	-100	56	N/A	4.9.3.: DETOXIFICATION: HEAVY METALS
Serum Related Protein Genes						
D-binding protein (DBP)	J03179	-3	OFF	57	N/A	2.14.1.: TRANSCRIPTION FACTORS
Transferrin	D38380	-4	OFF	58	N/A	6.3.: COAGULATION SYSTEM
Molecular Toxicology Marker Genes Set II (Drug Metabolism and Free Radical Scavenger Genes)						
Calmodulin (pRCM1)	X13933	+1.5	+2	59	N/A	2.13.: CALCIUM CASCADE
UDP glucuronosyl transferase, phenobarbital-inducible	M13506	+2	+2	60	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
UDP glucuronosyltransferase	U06273	+2	+2	61	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
UDP glucuronosyl transferase (UGT1.1)	U20551	+2	+3.5	62	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
Hydroxysteroid sulfotransferase	M31363	-9	-4	63	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
Estrogen sulfotransferase (Ste1)	U50204	+3	+3	64	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
CYP 2C13 (Cytochrome p450g)	M32277	+2	+7	65	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
CYP 2C23	X55446	+3	NEW	66	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
CYP 3A1 (Cytochrome p450/6 beta B)	AB008377	+3	+6	67	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
CYP 3A18	D38381/ X79991	+3	+10	68	N/A	4.2.4.: STEROID METABOLISM: EXCRETION
Cystathionine gamma lyase	D17370/ X53460	+4	+1.5	69	N/A	4.5.5.: SERINE FAMILY
Catechol-O-methyltransferase	M60753	-3	-1.5	70	N/A	4.8.3.: CATECHOLAMINE METABOLISM
Catalase	M25680	+2	+2	71	N/A	4.9.5.: DETOXIFICATION: HYDROCARBONS/ XENOBIOTICS
Glutathione S-transferase Yc2 subunit	S72506	+3	+3	72	N/A	4.9.5.: DETOXIFICATION: HYDROCARBONS/ XENOBIOTICS
Heme oxygenase	J02722	-2	-3	73	N/A	4.9.5.: DETOXIFICATION: HYDROCARBONS/ XENOBIOTICS
Unknown Function Genes						
Cytochrome p450 P49	M58041	+3	+2	74	N/A	9.: UNKNOWN FUNCTION
P8	AF014503	-2	-2	75	N/A	9.: UNKNOWN FUNCTION
Hematopoietic lineage switch 2 related protein	AF097723	-2	-2	76	N/A	9.: UNKNOWN FUNCTION
Non-Classified Genes						
40kDa ribosomal protein	D25224	-2	-2	77	N/A	1.2.1.: RIBOSOMAL PROTEINS
Ribosomal protein L18	M20156	-2	-2	78	N/A	1.2.1.: RIBOSOMAL PROTEINS
Ribosomal protein S7	X53377	+2	+2	79	N/A	1.2.1.: RIBOSOMAL PROTEINS
Nucleolar protein B23	J03969	-2	-2	80	N/A	1.2.4.: NUCLEOPROTEINS
Elongation factor 1-alpha	X63561	-2	-2	81	N/A	1.2.5.: TRANSLATION FACTORS
Proteasome RN3 subunit	L17127	+2	+2	82	N/A	1.5.1.: PROTEOLYSIS
26S proteasome, subunit p112	AJ006340	+2	+2	83	N/A	1.5.1.: PROTEOLYSIS
Chymotrypsin B (chyB)	K02298	+4	+2	84	N/A	1.5.1.: PROTEOLYSIS
Secreted phosphoprotein 24 precursor	U19485	-4	-4	85	N/A	1.5.1.1.: PROTEOLYSIS INHIBITOR
Voltage-gated Ca channel	AB018253	-2	-2	86	N/A	2.3.2.: ION CHANNELS
Acetyl CoA carboxylase 2 (ACC2)	AB004329	+11	+7	87	N/A	4.1.1.: FATTY ACID SYNTHESIS
D-3-D-2 trans-enoyl-CoA isomerase (dodecenoyl-coenzyme A delta-isomerase (DCI))	D00729/ M61112	+2	+3	88	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION

TABLE 1

Mitochondrial carnitine palmitoyltransferase II	J05470	+	+4	89	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION
Mitochondrial long-chain enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase alpha-subunit of mitochondrial trifunctional protein	D16478	+5	+4	90	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION
Long-chain 3-ketoacyl-CoA thiolase (last third of the mitochondrial trifunctional protein)	D16479	+2	+2	91	N/A	4.1.2.1.: MITOCHONDRIAL BETA OXIDATION
Trihydroxycoprostanoyl-CoA oxidase	X95189	-	-3	92	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
Peroxisomal 3-ketoacyl-CoA thiolase	J02749	NEW	NEW	93	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
Very-long-chain acyl-CoA synthetase	D85100	+2	+3	94	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
Carnitine octanoyltransferase	U26033	+3	+2	95	N/A	4.1.2.2.: PEROXISOMAL BETA OXIDATION
3-hydroxy-3-methylglutaryl CoA lyase	Y10054	+2	+2	96	N/A	4.1.3.: KETONE BODY METABOLISM
(Cytosolic) 3-HMG-CoA synthase	X52625	+5	+3	97	N/A	4.2.1.: CHOLESTEROL BIOSYNTHESIS
3-beta-hydroxysteroid dehydrogenase/delta-5-delta-4-ene-isomerase	M67465	+3	+3	98	N/A	4.2.1.: CHOLESTEROL BIOSYNTHESIS
Lanosterol 14-alpha demethylase	U17697/ AB004087	+3	+3	99	N/A	4.2.1.: CHOLESTEROL BIOSYNTHESIS
Lysosomal acid lipase	S81497	+3	+3	100	N/A	4.2.1.: CHOLESTEROL BIOSYNTHESIS
17 beta-hydroxysteroid dehydrogenase type IV (HSD IV) (peroxisome proliferator-inducible gene)	S83279	+2	+3	101	N/A	4.2.1.: CHOLESTEROL BIOSYNTHESIS
Fructose 1, 6 bisphosphatase	M86240	-2	-1.5	102	N/A	4.3.1.: GLYCOLYSIS/GLUCONEOGENESIS
Cytosolic malate dehydrogenase (MDH)	AF093773	+2	+2	103	N/A	4.4.1.: CITRIC ACID CYCLE
Cytosolic NADP-dependent isocitrate dehydrogenase	L35317	+6	+4	104	N/A	4.4.1.: CITRIC ACID CYCLE
Mitochondrial succinyl-CoA synthetase alpha subunit	J03621	+2	+4	105	N/A	4.4.1.: CITRIC ACID CYCLE
Mitochondrial cytochrome oxidase subunits I,II, III genes	J01435	+2	+21	106	N/A	4.4.2.: ELECTRON TRANSPORT CHAIN
Mitochondrial acetoacetyl-CoA thiolase	D13921	+2	+5	107	N/A	4.5.2.: ASPARTATE FAMILY
Carboxyl methyltransferase	M26686	+2	+4	108	N/A	4.5.2.: ASPARTATE FAMILY
Beta-alanine oxoglutarate aminotransferase	D87839	-2	-2	109	N/A	4.5.4.: PYRUVATE FAMILY
Serine dehydratase	Y00752	-3	-3	110	N/A	4.5.5.: SERINE FAMILY
Nucleus-encoded mitochondrial carbamyl phosphate synthetase 1	M12335	+3	+5	111	N/A	4.5.7.: UREA CYCLE
Arginase	J02720	-1.5	-2	112	N/A	4.5.7.: UREA CYCLE
Argininosuccinase (Argininosuccinate lyase)	D28501	-2	OFF	113	N/A	4.5.7.: UREA CYCLE
Argininosuccinate synthetase	M36708	-2	-6	114	N/A	4.5.7.: UREA CYCLE
Ornithine carbamoyltransferase	K00001	+2	+3	115	N/A	4.5.7.: UREA CYCLE
Ornithine aminotransferase	M11842	-2	-3	116	N/A	4.5.7.: UREA CYCLE
S-adenosylmethionine synthetase	X60822	-2	-2	117	N/A	4.5.8.: METHYL CYCLE
3 alpha-Hydroxysteroid dehydrogenase (3 alpha-HSD)	D17310	+2	+2	118	N/A	4.9.4.: DETOXIFICATION: ALDEHYDES
Apolipoprotein B	M27440	+4	+	119	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Sulfated glycoprotein 2	X13231	+1.5	+1.5	120	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Beta-globin	X06701	-2	-2	121	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Putative glycogen storage disease type 1b protein	AF080468	-4	-2	122	N/A	4.11.2.: PLASMA MEMBRANE SHUTTLING
Glucose transporter type 2 (GLUT 2)	L28135	-4	-2.5	123	N/A	4.11.2.: PLASMA MEMBRANE SHUTTLING
Sodium/bile acid cotransporter	M77479	+2	+4	124	N/A	4.11.2.: PLASMA MEMBRANE SHUTTLING

TABLE 1

Brain digoxin carrier protein	U88036	+2	+3	125	N/A	4.11.2.: PLASMA MEMBRANE SHUTTLING
Alpha-tocopherol transfer protein	D16339	+3	+7	126	N/A	4.11.3.: INTRACELLULAR STORAGE
Cholesterol 7-alpha-hydroxylase	J05460	+5	+2	127	N/A	4.: BASIC METABOLISM
Mitochondrial dicarboxylate carrier	AJ223355	-3	-3	128	N/A	5.6.:TISSUE ARCHITECTURE: MITOCHONDRIA
Laminin receptor	U04942	-2	-2	129	N/A	5.3.2.: INTERFACE WITH EXTRACELLULAR MATRIX
Mama	AF065438	-3	-3	130	N/A	9.: UNKNOWN FUNCTION
IRE-BP (Iron-responsive element-binding protein)	L23874	+3	+5	131	N/A	
Cytochrome P-450d methylcholanthrene-inducible gene	K02422	±1.0	±1.0	132	N/A	4.: BASIC METABOLISM
Tyrosine animotransferase	M18340	+3	-2	133	N/A	4.: BASIC METABOLISM
Acidic calponin	U06755	-2	+2	134	N/A	5.: TISSUE ARCHITECTURE
T-kininogen (T-KG)	M29083	±1.0	+2.5	135	N/A	1.: PROTEIN PRODUCTION
Trypsinogen II	V01274	N.D.	+4	136	N/A	1.: PROTEIN PRODUCTION
Ribosomal protein L23a	X65228	±1.0	+3	137	N/A	1.: PROTEIN PRODUCTION
Transmembrane protein rmp21.4.	X92097	±1.0	+2	138	N/A	4.: BASIC METABOLISM
Alpha-2u globulin	U31287	±1.0	-3	139	N/A	4.: BASIC METABOLISM
Submaxillary gland alpha-2u globulin (distinct gene)	J00738	±1.0	-3	140	N/A	4.: BASIC METABOLISM
Fetuin (alpha-2-HS glycoprotein (AHSG))	X63446	±1.0	-3	141	N/A	4.: BASIC METABOLISM
60 kDa protein and non-specific lipid transfer protein	M62763	±1.0	+4	142	N/A	4.: BASIC METABOLISM
Lecithin:cholesterol acyltransferase	U62803	±1.0	-2	143	N/A	4.: BASIC METABOLISM
Pancreatic amylase	J00703	±1.0	-4	144	N/A	4.: BASIC METABOLISM
Mevalonate pyrophosphate decarboxylase	U53706	±1.0	+2	145	N/A	4.: BASIC METABOLISM
Branched chain alpha-keto acid dehydrogenase E-subunit	M94040	±1.0	+1.5	146	N/A	4.: BASIC METABOLISM
Carboxylesterase	M20629	±1.0	+1.5	147	N/A	4.: BASIC METABOLISM
Apolipoprotein A-I	X00558/ M00001/ J02597	0	-4	148	N/A	4.11.1.: EXTRACELLULAR TRANSPORT
Glucokinase regulatory protein (GCKR)	X68497	±1.0	-1.5	149	N/A	4.: BASIC METABOLISM
Long chain acyl-CoA dehydrogenase (LCAD)	J05029	±1.0	+2	150	N/A	4.: BASIC METABOLISM
Glutathione S-transferase	M25981	±1.0	+5	151	N/A	4.: BASIC METABOLISM
Alpha albumin	X76456	0	OFF	152	N/A	4.: BASIC METABOLISM
Liver fatty acid binding protein	J00732/ V01235	±1.0	+2	153	N/A	4.: BASIC METABOLISM
HBP23 (heme-binding protein 23 kDa)	D30035	±1.0	+2.5	154	N/A	4.: BASIC METABOLISM
CYP 3A2 (Testosterone 6 beta hydroxylase)	U09725	±1.0	+8	155	N/A	4.: BASIC METABOLISM
Core protein heparan sulfate proteoglycan (HSPG) (Syndecan2 (SDC))	M81687	±1.0	-2	156	N/A	5.: TISSUE ARCHITECTURE
Alpha-fibrinogen	X86561	0	-4	157	N/A	6.: EXTRACELLULAR ENVIRONMENTAL REGULATION
EST AA893518		±1.0	+20	158	N/A	9.: UNKNOWN FUNCTION
Cysteine-rich protein 2 (CRP2)	D17512	±1.0	-2	159	N/A	9.: UNKNOWN FUNCTION
Glu-Pro Dipeptide Repeat	U40628	0	-8	160	N/A	9.: UNKNOWN FUNCTION
Preprocarboxypeptidase A2 (proCPA2)	[P19222-protein]	±1.0	-2	161	N/A	9.: UNKNOWN FUNCTION
Alpha-1-macroglobulin	M77183	+5	-1.5/ ±1.0	162	N/A	1.: PROTEIN PRODUCTION
Ribosomal protein S24	M89646	-2	±1.0	163	N/A	1.: PROTEIN PRODUCTION
Cytochrome p450 isozyme CMF1b	M22329	-2	±1.0	164	N/A	4.: BASIC METABOLISM
Epoxide hydrolase	M26125	+1.5	±1.0	165	N/A	4.: BASIC METABOLISM
(Short-chain) enoyl-CoA hydratase (ECHS, SCEH)	X15958	+2	±1.0	166	N/A	4.: BASIC METABOLISM
Stearoyl-CoA desaturase	J02585	+5	0	167	N/A	4.: BASIC METABOLISM
Aquaporin (9)	AB013112	-3	±1.0	168	N/A	7.3.1.: PORINS
Hepatic tumor cytochrome P450 (CYP4F1)	M94548	+3	±1.0	169	N/A	9.: UNKNOWN FUNCTION

TABLE 1

N.D. = not determined						
NEW = *de novo* synthesis, trancription was initiated in response to troglitazone						
OFF = transcription was turned off in response to troglitazone						

TRADOCS:1319690.1(S@@201!.DOC)

Below follows additional discussion of nucleic acid sequences whose expression is differentially regulated in the presence of troglitazone.

HEPATOX1

HEPATOX1 is a novel 185 bp gene fragment. The nucleic acid has the following sequence:

```
 1 actagtcatc ctttcccttt accagtacgg agcagcacaa cactagagtg gggagggaca
61 aaaagtttca aggtaggact ggaccatcct attccaatag ccaacaccct caggggttgt
121 tctctgtgcc tgaccctagc ttggagttgg gtgtgggaaa ggcatgggag cattcttcat
181 catga (SEQ ID NO:1).
```

HEPATOX2

HEPATOX2 is a novel 56 bp gene fragment. The nucleic acid has the following sequence:

```
1 ggatcctcca gtcacactct ccaccttgcc ttccagctgt gggaacagtt tcatga (SEQ
ID NO:2).
```

HEPATOX3

HEPATOX3 is a novel 572 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
1 actagtactg ggggagggaa gggggtagaa ttggctcaat tg (SEQ ID NO:3).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1     cggccgcaaaggttttgttttgtcttataaaggtaaagcttgtcggcacgtgaagcgagaagtccattagttgacatttt
81    cctgtttttagaagctactccctttttggggagcaggtctgcatgacttgcttacctgatgtgtccctgttgtcaggtcag
161   agtgtgtttctgtgagctgtcattgtggggaaccaattgctttgtcatatatctggttccgaactataacttgtttggga
241   agtctactgatgttcctgctggttttcccaagtgtgctttcgcaatgcatggtatccacgtcacttggaaagtgggcttt
321   gttgtctgtctgaatttagggtcacttaagaaaggggcagttaaaaataaagcacaatgcctcacatgggacaaagttcc
401   aaaatgccaaattcttctatttttacaacaacctgtagttatgtaaaataactagtactgggggagggaaggggtaga
481   attggctcaattggaaagactctcccacttcacactaagcttgtttgtcaccatgagatagcattggctgcccaggatgc
```

561 tgctatttaaaa (SEQ ID NO:4).

HEPATOX4

HEPATOX4 is a novel 1770 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 gggccctatg tgcatctcgg gaaaatactt gtgttcattg tgcagagttc attcctcatc
 61 tgaacgttat gtcctaaagg agtgcacatg gctaaagtgt agctaatgat atggcagaac
121 tgacagaggg ttcctcttga acatggaact ggtaagccat gtctggacag tggggaaatt
181 catgttgcta ggatgactgg aggggcagag aggctctcgg ggtgaggtct gcagctgtgt
241 ggagacatcc tgagagcccc ccactgtggg tgcatctccc agctgctcca tgggaataca
301 tttcttgatc ggggaacacc ccttctgcat atgttagtgt ctaga (SEQ ID NO:5).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1    tgccaccgagactaccgaagaattcgcagccctttttttttttttttttcagttttaaaagaaggttttgtttatttttt
81   aatatgtgtattacaggaaacaaaacacaagagggcaagaacaaagtcacccagtcacaagcagctcagctcggtgtgtc
161  gttctagatcttgcaggtgtacatacacttgggtaatggagattatacagtatgtaccatgcttccccacattgtgaata
241  ttcaccaactaaaaaccacgtgccacacgaggatttgatagccaagaatacaccatgccatctcaatctgtcagacaaaa
321  ttgtaatcttgtcttccttggcaaaagaaacttcatagaagacaaaaacggcagcaggcctctagatagacgtactggca
401  aagttacaactgaattctgcattcgctcagtgctcagctcacaaggaaacacaaagcaacagaccacactaagtctagca
481  ttcctaagctcatcatcaagatctagacactaacatatgcagaaggggtgttccccgatcaagaaatgtattcccatgga
561  gcagctgggagatgcacccacagtgggggctctcaggatgtctccacacagctgcagacctcaccccgagagcctctct
641  gcccctccagtcatcctagcaacatgaatttccccactgtccagacatggcttaccagttccatgttcaagaggaaccct
721  ctgtcagttctgccatatcattagctacactttagccatgtgcactcctttaggacataacgttcagatgaggaatgaac
801  tctgcacaatgaacacaagtattttcccgagatgcacatagggcccctgggtacagacagctgtcttctggatgtagtag
881  tactacagtctctcccctctcccatcctcttctctgcacaactcctaagcgcagatgatgctagtgagaccataacaaaga
961  ccagcctgacttctagtgactcttgctactgctaacatcctttaggcactgggactatttctaatgcctggcaactggtt
1041 aactgaaatgcaaatctaaccggttctagacagcttgagccaacagctacagtgccctttctcagcactagtctttgcc
1121 ctagcccacctccaccacctggtaacagcatctgctcaaaggccaactttgccctccacttccccggtgtgaccattcct
1201 atgaaataaaacaatctaaaaagtggtgttaccatggtgtgcagttttacatttacacattactgtatgcatgaggctaa
1281 aagacaagttgtgaatgacttctcctatcatgtggatcctggaaatcaagctcaggtcacaagcttggtgacaagcacct
1361 ttccatgccgagccaccctgccatttcatcatgcacactttcttactgatttgagacaaagttcatagatagcactggct
1441 gaaccggaacacagagatccacctgcctcagcctagaattaaaggtgtgctctaccacaccaagactattaatgcacctc
```

1521 tgaagactatctagagactgatgttcccaatacgcacttccgtttccccagcatgagtacactactgtggatacactgta

1601 cacctaacaaaggctgcagtcaaaaactgccttctcaacacccaggtttccggtctagcccctcttcctgttctctccac

1681 ccaggggtcacgtgtgcaactgccatgtctttgtgcagagatgaacattttcatttcctaaatggttttatgagtttttt

1761 tctacaaatn (SEQ ID NO:6).

HEPATO5

HEPATOX5 is a novel 755 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 gggccctgct gggctagact ggcaaaggtt agaaagtgac accagctcaa gagaaccact

61 cagacttact ggagaaagct gtaagttgcc ctttgatgta attgggactc actctaga (SEQ ID NO:7).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 ttttttttttttttttttttgtgctttaaatagattttatttgtttatattatttttttctcccatacaggaatagcat

81 tacaacctaacattcagaatcctgttacacacatacacaggcatgccacatgaccaggtcgaggtggttgtgtccttgag

161 tctgtcgacacatcacaacatggtcagtcaaagagttttcacttcagccagtctcaacacaaaacacccaacagagatgc

241 actcaacttgttggttccatttggaactaggtggtagggaaagaaggtatttgaagcaggctgtggtttgtctgcattca

321 gtcccctcacacaaagctacgtggatcttgaaagccacccaggagctctggtagggtccttgttgcacctaagacattat

401 aggtcccagcaagttgctgagagggttccaggctggggagcaggcatactccaaaacagcaccaaactgcactcacacac

481 acagggccctgctgggctagactggcaaaggttagaaagtgacaccagctcaagagaaccactcagacttactggagaaa

561 gctgtaagttgccctttgatgtaattgggactcactctagaaacaggtccctctgacatccttagttttggcaaagtgct

641 ggtagtactgacattaatacaaactggtacaaatttctctagggtctctgaaatcaccaatcaagacagactctctaacc

721 tcctcctgtctagggttgcataggactntgtgaat (SEQ ID NO:8).

HEPATOX6

HEPATOX6 is a novel 675 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 gaattcttct actaccatgg cctgtcttcc ctctgcaacc tggggtgtct gctgagcaat

61 ggtctctgcc ttgtgggcct cgccctgggg ctcaggagcc tctaacatgg gccc (SEQ ID NO:9).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1     nctggggctcttggtcccttcacctgaccttccggtgctccaacgggcggcagaatggaagaaggtgaggacccaggaag
81    tctgattaaagtgatccacttgctggtcttgtctggtgtctggggcatgcagatgtgggtgacctttgcctcaggcttcc
161   tgcttttccggagcctcccgaggcacacgtttggacttgtgcagagcaagctcttcccagtctattttcacgtctccttg
241   ggttgtgccttcatcaacctctgcatcttggcaccacagcgggcctggatcaacctcacgctgtgggaaatcagtcagct
321   taccctactgcttctgagtctcacactggctaccatcaatgctcgctggctcgaggctcgcaccacagctaccatgtggg
401   ccctgcagagtatagagaaagagcgaggcctggggacagaggtgccaggcagcctccagggccctgacccctaccgccag
481   ctgcgggagaaggacccccaagtacagtgctctccggcaggaattcttctactaccatggcctgtcttccctctgcaacct
561   ggggtgtctgctgagcaatggtctctgccttgtgggcctcgccctggggctcaggagcctctaacatgggccctgtgttt
641   caataaatgtttctttggaaaaaaaaaaaaaaaaa (SEQ ID NO:10).
```

HEPATOX7

HEPATOX7 is a novel 874 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 caattgcaac cataatgaag atgcaggagt ggaatgccgc tgacctggga gcctgagaag
 61 tcatccgtgt gttcccaggt gtctttgggc accacccaca tggagatct (SEQ ID
NO:11),

  1 caattgtgag agccccagga gctcttcctg cagtcccaca aagtggcctc tgtccccaga
 61 caactcacat catccagcca gatcttccca gtgccacctc caaagctgcc aagtcccctc
121 ccgctggagt aaccgagcat gcggcagaag acagtggcat cgttattatc ccaaccatca
181 tcgcaaattg tcccccagac attgttatag taaacttcag ctcggcctct gctggtgccg
241 cccacgatcc gga (SEQ ID NO:12),

  1 agatctggct ggatgatgtg agttgtctg gggacagagg ccactttgtg ggactgcagg
 61 aagagctcct ggggctctca caattg (SEQ ID NO:13).and

  1 ggatccccag ggattgccgg taatccaggg gctgcaggtt tgaaaggaag caaggggac
 61 acaggaattc aaggacagaa aggcacaaaa ggagaatcag gagtcccagg tcttgtaggc
121 agaaagggag acactggaag ccctgggctg gcaggcccca aaggagaacc tggacgagcc
181 ggtctgaagg gagaacctgg gatgaaaggg tcttctggcc agcaaggaca aaagggagaa
```

241 aaaggtcaaa aaggcgattt caacctggtt gtccgga (SEQ ID NO:14).

The cloned sequences were assembled into a contig resulting in the following consensus sequence:

1 agatctccatgtgggtggtgcccaaagacacctgggaacacacggatgacttctcaggctcccaggtcagcggcattcca

81 ctcctgcatcttcattatggttgcaattgtgagagccccaggagctcttcctgcagtcccacaaagtggcctctgtcccc

161 agacaactcacatcatccagccagatcttcccagtgccacctccaaagctgccaagtcccctcccgctggagtaaccgag

241 catgcggcagaagacagtggcatcgttattatcccaaccatcatcgcaaattgtcccccagacattgttatagtaaactt

321 cagctcggcctctgctggtgccgcccacgatccggacaaccaggttgaaatcgccttttgacctttttctccctttgt

401 ccttgctggccagaagacccctttcatcccaggttctcccttcagaccggctcgtccaggttctcctttggggcctgccag

481 cccagggcttccagtgtctccctttctgcctacaagacctgggactcctgattctccttttgtgcctttctgtccttgaa

561 ttcctgtgtccccctttgcttcctttcaaacctgcagcccctggattaccggcaatccctggggatcctgttggcccagga

641 ggaccagtgcgccccggctcacccttggcacctgattgtcctggtgcaccaggagggcctggaggaccctgcactccaag

721 ttttccttgatcacctttggctccaggagatccagccaaacctggtaggcctggttttccagcatcgcctttatttccct

801 gggctccaggggaccccgtgacccccctgtctcccttcatgcccgtgtcccctttgcttcctggaaggcctagg (SEQ ID NO:15).

HEPATOX8

HEPATOX8 is a novel 546 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 gctagcatcc ttccactgct tctcccatgg ctccggagcc ctctgtgcat gcaacttgct

61 agccaagctt (SEQ ID NO:16).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 aagcttggctagcaagttgcatgcacagagggctccggagccatgggagaagcagtggaaggatgctggcataccctggg

81 gcaggccacctcatagagccgccttattctccctttgctttgcgtgctgggattctgtcttgggcaaacccatgctttg

161 gggaggagatcccatcgctcatgcagcagctcaggtgcactcctggagagagatccaaaagttcttcaggcaacacctcc

241 ttcagtctggaggcaaactctaaatagtgtgaccatgtaggacagagtaaagggcagggagtgaattagagaagaggttg

321 gaattggtctggggattaaagcatcatcagttaggggctgcagggcaagggtggggatgggggacagaattctcctggct

401 ttcttactctcacaatcattgcaattcaattggaaggaaaccagacaaattcacagacccggacacacgtgcacctgagg

481 tgtttttggtgaatttgtatgtagtttgaaattttctaaataaattttaaattttaccatgcaccn (SEQ ID NO:17).

HEPATOX9

HEPATOX9 is a novel 407 bp gene fragment. The nucleic acid has the following sequence:

```
  1 caattgaatt gcgatgattg tgagagtaag aaagccaccc ttgccctgca acccctaacg
 61 atgatgcttt aatccccaga ccaattccaa cctcttctct aacgcactcc ctgcccttta
121 ctctgtcctg catggtcaca ctatttagag tttgcctcca gactgaagga ggtgttgcct
181 gaagaacttt tggatctctc tccaggagtg cacctgagct gctgcatgag cgatgggatc
241 tcctccccaa agcaagggtt tgcccaagac agagtgccag gacgcaaagc aaaagggaga
301 ataaggcggc tctataaggt ggcctgcccc agggtatgct aacatccttc cactgcttct
361 cccatggctc cggagccctc tgtgcatgca acttgctagc caagctt (SEQ ID NO:18).
```

HEPATOX10

HEPATOX10 is a novel 1149 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 actagtagca catctgtaat accaaactac aggacaagtc ttaacaagag gtttgtgttc
 61 ttgaacgtag cacttgtcta ccaggcactg tagaagagaa tgaggaaaag ccaggacctg
121 ctcanggagc ttaaggggtt gggntnntgg gtgggatann tggacagtaa ccacttccta
181 agcaacctgg gttttctccc cagggaaggc aaacatgcaa agctt (SEQ ID NO:19).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1   ttttttttttttttttttatagctagtatgagttagcattttaattttaaatcctatataaaataatatatttttttttttt
81  ttttttctttatgaaatttatttttcttatataaattatgtatttctctgggcagacagccttcaccttattgcactagta
161 gcacatctgtaataccaaactacaggacaagtcttaacaagaggtttgtgttcttgaacgtagcacttgtctaccaggca
241 ctgtagaagagaatgaggaaaagccaggacctgctcaggagcttaagggttgggttgggtgggatatggacagtaacact
321 tctaacaactggttttctcccaggaaggcaaacatgcaaagcttccatctggattttgggagcagggacaggttggaagg
401 caactctgctccctttgtttctgagagaagtgggcacttacctcgcatttgtaacaaggcctaactgatgctttggggca
```

481 acaatcaggtgtggagttttcatcaccagatttataattagcagcaattaagaatgggccaccatcactgccacagtgag

561 gtgacccatcctggcaccagggacccagaagctgtattctagataggtgccactaagatccagagctgggacctcaaact

641 gctccttatctcaccatggctctgcaacactagttagcaacactccaggtacagggttcctcaaccactggtcaaagcta

721 tgatatggatggagtctgtaggcagagtgaaggggatgcctgagtaggaagagggaaggtcactgggcagtgattgtaac

801 atgattagaccctccttgcaggaggtaggggtgaacaaagggttgggttagaaggagctgcaagtctccaaagctcccct

881 cactctgcctttgtactacactgaagacaagttgctcacatactctaaagcacattcttgatacaggaatagggacttgg

961 gcaagcggcgatttggtattaggcaagagccacattaagccttcatgaagggcagccccccacagccttggcattgagac

1041 tgcgtaggtcacagtaggtctggccagtacccacaaagacgatgggtttgcttgtgatgtatgtcatagaaatagcagct

1121 cccaccttgtcatcaatggtgtcaaattt (SEQ ID NO:20).

HEPATOX11

HEPATOX11 is a nove572 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 gggccctgaa catcatcaat gaagagagaa cagcacttca aaggaggccg ttgcaacagt

61 cctgtctccc tgaccctgag gaaggacaac ttttatatgc aaatatgtac a (SEQ ID NO:21).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 ccatggaatgcacacgacggttctgccctttatactaatcgagatgagaacgtcccaccaccagtaccccagcagaagca

81 gtgggctcgccgccatatgcaccttctccgtgggctatatattgtgggtgtgctggatacatcatgtcaccggcatgtgg

161 gtgtacccttttcctggaacacattggctcaggagccaggatcatcttctttgggtcgacaaccctcctaatgaattccct

241 gtacttgctcggagaagcactcaacagctacatctgggatgccccgagaagtttagaagaagaaaaagagaagcctaaac

321 tggagtgagatccaagtctgaggagaagcggggaggagctgcccttgaattctttgggccctgaacatcatcaatgaaga

401 gagaacagcacttcaaaggaggccgttgcaacagtcctgtctccctgaccctgaggaaggacaacttttatatgcaaata

481 tgtacaaaatataggatatgttgcctttgaaaataactcatactccttgttgtactctacccaagttcattattgatgacat

561 tttttcctcttttctagttttaaaattaaagttggtcattgagttttcatttctgcaataataactctacttcgtcaagt

641 aaagaaaaaacatcaaatatatttcttattgaggaatatgataagattggaggggcagatgtgggtggaggcattcatag

721 ttagtcttaaaggaggctggggtcatgacactagggtacaaaatggacccaggagaatgcagggtaggtggttccggtgg

801 aggtagaatcaggcttgcctccgtgagcagtgttagtaatagtatccgtttagtcaaaaccctgagtgtgcttggatctc

881 cctgtttgatctgtgaaactcttcccatatacaaacgtgactgttaataaaaaccaagcaacccaaaaaaaaaaaaaaaa

961 a (SEQ ID NO:22).

HEPATOX12

HEPATOX12 is a novel 667 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 actagtgaag gcgtttaaga agaaatttgc ctgcaatggt actgtaattg agcatccaga
 61 atatggagaa gtaattcagc tacagggtga ccagcgcaag aacatatgcc agttcctgat
121 agagattgga ctggctaagg acgatcagct gaaggttcat gggttttaag tgcttgtggc
181 tctctgaagc tt (SEQ ID NO:23).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1     ttttttttttttttttttatggctcttttcagtttattgcatgaaggagttacactagtccaagttaaaagcggacccc
81    aaatgattacattatacaagctgtgaggtttttaaacttgtgacaagggacagaagggaaattctactcattgcaaggaa
161   atcctcacttaagcttcagagagccacaagcacttaaaacccatgaaccttcagctgatcgtccttagccagtccaatct
241   ctatcaggaactggcatatgttcttgcgctggtcaccctgtagctgaattacttctccatattctggatgctcaattaca
321   gtaccattgcaggcaaatttcttcttaaacgccttcactagtttctttttatcgtaatcatcagcgatcccttggacagt
401   tgtaagggtcttcctgccgtttctctgttgaattcttatatggatataatcctcagtgccagcaggaagcaggtcatcac
481   ccttacttgcatcagcaaaggggtcgaaagagtggaggttctggatagcggacatacgatacgattccttttcctcggtg
561   gaaacggcctgcggaaggcggctgcgggagaaggcgggcgggggggacggagcgtcgggaagcgaggggctcgaggggg
641   aggcagctgagtcctcggcggcggctc (SEQ ID NO:24).
```

HEPATOX13

HEPATOX13 is a novel 580 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 caattgatgc tgatgtgaca gtgataggtt ctggtcctgg aggatatgtt gctgccatca
 61 aagctgccca gttaggcttt aagacagtct gcattgagaa gaatgaaaca ctaggaggaa
121 catgcttgaa tgttggttgt attccttcaa aggctttatt aaataattct cattattacc
181 atttggccca tggaaaagat tttgcatctt ggggaattga ataccagaa gttcgcttga
241 atttagagaa gatgatggag cagaagcgtt ctgcagtaaa agcattaaca gggggaattg
301 cccacttatt caaacaaaat atggttgttc atgtcaatgg atttggaaat ataactggca
```

361 agaatcaggt tacagctaca acggccgatg gcagcactca ggttattggt acc (SEQ ID NO:25).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 ggtaccaataacctgagtgctgccatcggccgttgtagctgtaacctgattcttgccagttatctttccaaatccattga

81 catgaacaaccttattttgtttgaataagtgggcaattcccccctgttaatgcttttactgcagaacgcttctgctccatc

161 atcttctctaaattcaagcgaacttctggtatttcaattcccctagatgcaaaatcttttccatgggccaaatggtaata

241 atgagaattatttaataaagcctttgaaggaatacaaccaacattcaagcatgttcctcctagtgtttcattcttctcaa

321 tgcagactgtcttaaagcctaactgggcagctttgatggcagcaacatatcctccaggaccagaacctatcactgtcaca

401 tcagcatcaattggttgatctgaataagttcttaatggaactgaagaagccccctgtaggccatgagacagccgattgaa

481 atggcctttcttggccaaggagcagtacacacgactccagctctgcatctcccctcgttcccgtccctacaccagggctt

561 tcacctccgccctcgtgccg (SEQ ID NO:26).

HEPATOX14

HEPATOX3 is a novel 327 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 tgtacaagaa cagcgtcagt acagctcttg agcacacact cagtgccttg ctgactccag

61 ctgaacagac aagcagctgg aaattcaaga ccagcctgaa tgacaaagta tacagccagg

121 aatttgaagc ctacaacact aaagacaaaa ttggtatcga gcttagtgga cgggctgacc

181 tctctgggct gtactctcca attaaagtgc cgtttttcta cagtgagcct gtcaatgttc

241 ttaatagctt ggagataaat gatgcctttg acgagccccg agaattcaca attg (SEQ ID NO:27).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 tgtacaagaacagcgtcagtacagctcttgagcacacactcagtgccttgctgactccagctgaacagacaagcagctgg

81 aaattcaagaccagcctgaatgacaaagtatacagccaggaatttgaagcctacaacactaaagacaaaattggtatcga

161 gcttagtggacgggctgacctctctgggctgtactctccaattaaagtgccgtttttctacagtgagcctgtcaatgttc

241 ttaatagcttggagataaatgatgcctttgacgagccccgagaattcacaattgtcggggaaaaggttgaaaagtgctaa

321 gccaatt (SEQ ID NO:28).

HEPATOX15

HEPATOX15 is a novel 650 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 tgtacactct ggtgaaatgg ttctcacagc aggagcatca cagccagact ggacattctc
 61 tcaaaggggt acgagttgca gttctgaagg cccctggctt tggttgttca caaagttcag
121 tcctgtttac tgtgatcctt ggtgtcctcc tcgtctgtgt actctgatgg ctcttccccc
181 ggctttagga gtctgcctac ataatcatat ttttctttaa actgcatttc ccactctcga
241 acgctctcca tctgcactgc gttcaaatct gagaggtcat cgtactcatc tctaagtgca
301 tccttatcca ggcagaaggt cgccagtccc ctggaggcgt ccctgccagc aaagatgcca
361 tatgggccc (SEQ ID NO:29).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1   tttttttttttttttttgagatgctcgaagtttattgcaaagaggaaggcggggttggtgtagggaggtcaaggagaaaa
81  ggaagaggaggaaggaagggaggaacatcgagaggagaggagggtaaaataacccggagactttcttgctgttgagaagg
161 tcctgtctccttttcagggtgatgaagcccaccagacatcacaaacaactgcaacaggttcaccggcaggcagcacaggc
241 aatgcctcatattcagatcttcacagttgggcatagtatcttgtacactctggtgaaatggttctcacagcaggagcatc
321 acagccagactggacattctctcaaaggggtacgagttgcagttctgaaggcccctggctttggttgttcacaaagttca
401 gtcctgtttactgtgatccttgctgtcctcctcgtctgtgtactctgatggctcttcccctggctttaggagtctgccta
481 cataatcatatttttctttaaactgcatttcccactctcgaacgctctccatctgcactgcgttcaaatctgagaggtca
561 tcgtactcatctctaagtgcatccttatccaggcagaaggtcgccagtcccctggaggcgtccctgccagcaaagatgcc
641 atatgggccc (SEQ ID NO:30).
```

HEPATOX16

HEPATOX16 is a novel 499 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 tctagatgtg tcttgcacag ccatggctcc agcttcacca ggatcaggct cttctccttg
 61 ggcctccctg ctgacaagtc ttgcccaaaa cccaggtaga tggtgtaatg tggtgatcct
121 tgtcgcctcc gggccc (SEQ ID NO:31).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1     ggtaacccaaccaattctttattgtaaagatgagtgctacaaaggaggcctagtgtgtagggagtgggggtcaccttct
81    ttcaagggactccacaaggtccactagagatgacatatagccaaggaataagcctgagccagggcagcagtggttctgaa
161   ctctattggaagttggggttctgaatcaagcccactgacccaggtccatgaggaagtgttcaatgtcatcgtagagacta
241   ttggtgctagacaagcacaggccaagactgctgctgtccagagaggacacaccctcacgctgcacgccttctagatgtgt
321   cttgcacagccatggctccagcttcaccaggatcaggctctttctccttgggcctccctgctgacaagtcttgcccaaaac
401   ccaggtagatggtgtaatgtggtgatccttgtcgcctccgggcccggaactcctccagttctcggaagaaatgctgaagt
481   caaagatgggagtatggca (SEQ ID NO:32).
```

HEPATOX17

HEPATOX17 is a novel 315 bp gene fragment. The nucleic acid has the following sequence:

```
  1 gaattctgaa aataatagtt ttctttctgt gcgttcttac atacatcaag cccctccagt
 61 tcctgacagt gaagatacag cttccccgtg tcttcatgag acacatcatg ctctcttctg
121 tggaatagtc ttccagggcc acgaaatgac acctgttgaa ctctttcatt ccgactgagt
181 tttctccaaa atccagtcca catacttggc cacgttggtg tagacacccg gcctctcttt
241 ctggccgcag ccttcacccc agcttgtgac gcccaccaag tgccagaccc cattgtgttt
301 gcaggacagg ggccc (SEQ ID NO:33).
```

HEPATOX18

HEPATOX18 is a novel 760 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
  1 ccatggaaaa cgttttcaga agctacacga atcgctcttt caacagacca aatgtcacca
 61 cggtgacaag caaatgtcac caaatgtcac cacggtgaca agcagactct tgcatgcttt
121 ctgtatcact ctcgatgcac cgcgtgctct gactgtcttt cctcacagga gagtaagcac
181 agggaacggg ctgcagtgtt ttcctctgtg gaagatcaac tggcaaccct ggccttttta
241 aaaccgagag tagcagaatt aatagaggaa ctaattgggg taaataataa aggttttaat
```

301 ccagatctag a (SEQ ID NO:34).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 ttttttttttttttttgggttttttctggtttattaaaacaaaactccatttccatgcttctcaggttgtaccatcat
81 tgggcataaaaatcagtcaacatttgcaaggcagtgtcactcacagaatcacccagtgctattttacagtggtcattctc
161 tttacaacaatcaggcattgctttgtacaatgggcaaacgctataacaaacgccttccataaacagttagccctttagct
241 attgggaaggtcaagagtcgcttgaagcagtcactggtcgcactaaacaaagtaccaggtggcagctctcaagcccctaa
321 aaagtgtttttaaaaaacagatttacatcccatagtaagtcttagaaacttcaaagatagggcctatcatttcaaaacac
401 tacggttatctacggataaaaatctgcatccaactgtgcagtgttttcttctagatctggattaaaacctttattattta
481 ccccaattagttcctctattaattctgctactctcggttttaaaaaggccagggttgccagttgatcttccacagaggaa
561 aacactgcagcccgttccctgtgcttactctcctgtgaggaaagacagtcagagcacgcggtgcatcgagagtgatacag
641 aaagcatgcaagagtctgcttgtcaccgtggtgacatttggtgacatttgcttgtcaccgtggtgacatttggtctgttg
721 aaagagcgattcgtgtagcttctgaaaacgttttccatgg (SEQ ID NO:35).

HEPATOX19

HEPATOX19 is a novel 499 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 ggatcctttt gcggcattcc acattaacaa gggccttgtg agaaagtata tgaactctct
61 tctgattgga gagctagctc cgga (SEQ ID NO:36).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 tgtacaacatcagcgacttcagtcgccgccacccgggaggctcccgggtcatcagccactacgctggtcaggatgccacg
81 gatccttttgtggcattccacattaacaagggccttgtgagaaagtatatgaactctcttctgattggagagctagctcc
161 ggagcagcccagctttgaacccaccaagaataaggcgctcactgatgaattccgggagctgcgggccacagtggagcgaa
241 tgggcctcatgaaagccaaccatctcttcttcctgttctatctgctgcacatcctgctgctggacgtggccgcctggctc
321 actctttggatctttggaacttccttggtgcccttcaccctctgtgcagtgctgctcagtacagttcaggcccaggcagg
401 ttggctacagcatgactttgggcacctgtccgtcttcagcacctcaacatggaatcacctggtacatcattttgtcattg
481 gccacctgaaggggccc (SEQ ID NO:37).

HEPATOX20

HEPATOX20 is a novel 1083 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

```
   1 actagtagga taattattga ggctgtagct tgggttagaa aatattttgt tgctgcttca
  61 gttgatcgtg ggctttttt gttggctaga agtgggatga tagctaaaag gcttatttcc
 121 aatcctactc atattaggag taagttggag ctagatatag tgattgcagg ccctatgagg
 181 atagtaaagt aaataatgat tagggtgatt gggtttatta gtacgggaag ggtttagacc
 241 aacattttcg gggtatgggc cc (SEQ ID NO:38).
```

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

```
1     atgtgattagttgggaggaaagcggtagggtaagggtacttaatactgtgagggttgggaggataaaattatgttttggg
81    ttttggtggtggtgggagattattttggagttgttgttggttgggaatatggtgagggatatagagtaaattagtcgtgt
161   atagaagaataggcttaatagggctatgatagctattagtgttgatagggttgagcagttgtttttagaagctcggaga
241   taattgttcattttggtaaaaatcctgtgagagggggagtcctcctagtgataggaggatgatggatgctatggttagg
321   attatgggagttttatttcatgcgagtgagagtgtgttgattgttgttgctgagtttgttatgagtgtgatgaatattgg
401   aacagtaagtaggatgtaaattgttaagtttaggagggttaagttagggttgtatggaaggattgctgttattcatccta
481   tgtgggcaattgatgaatatgctatgatttttcgtgtttgggtctggttaagtcctcctcaggcgccaacaaagactgat
561   gaaattgcgagaatggtggtgatagttgggtttaggagttgataaaattggtatagaattgatagtggagcaatttttg
641   tcatgttagtaagattaatccaatgtgtaggggaattccttgggtgacttcgggtagtcagtagtggaatggggctagtc
721   caagttttatggctagtgaaatgagtattatgttgagtagtatgttattggtttgttgttggagggttcatattcctgat
801   tgtttgtagttgaggatgatgactagtaggataattattgaggctgtagcttgggttagaaaatattttgttgctgcttc
881   agttgatcgtgggctttttttgttggctagaagtgggatgatagctaaaaggcttatttccaatcctactcatattagga
961   gtaagttggagctagatatagtgattacaggccctatgaggatagtaaagtaaataatgattagggtgattgggtttatt
1041  agtacgggaagggtttagaccaacattttcggggtatgggccc (SEQ ID NO:39).
```

HEPATOX21

HEPATOX21 is a novel 580 bp gene fragment. The nucleic acid was initially identified in a cloned fragment having the following sequence:

1 tctagatttg gatggttgct gggtgataca agttgatgag agaatactgc caacatcata

61 gcagaatgaa attcttcctg ctgctttccc tcattgggct ctgctgggcc c (SEQ ID NO:40).

The cloned sequence was assembled into a contig resulting in the following consensus sequence:

1 gaattcatcctcatttccagacctggagcatattttgtagctaattggttgatatctttcccaccatggtctaaaaggac

81 tgttgaccacaatattttcattgggtggagaaacctgcacgcctccaaatccattaggagctaaatatcgctcacactcc

161 ttagcaatatcaacccagcgccactcgaacaaatggacaatagaagaccgtccataaaaagtatgtgggtcatattgggc

241 ccagcagagcccaatgagggaaagcagcaggaagaatttcattctgctatgatgttggcagtattctctcatcaacttgt

321 atcaaccagcaaccatccaaatctagactgcatgggatattgcacacgcgggaaag (SEQ ID NO:41).

GENERAL METHODS

The HEPATOX nucleic acids and encoded polypeptides can be identified using the information provide above. In some embodiments, the HEPATOX nucleic acids and polypeptide correspond to nucleic acids or polypeptides which include the various sequences (referenced by SEQ ID NOs) disclosed for each HEPATOX polypeptide.

In its various aspects and embodiments, the invention includes providing a test cell population which includes at least one cell that is capable of expressing one or more of the sequences HEPATOX 1-169. By "capable of expressing" is meant that the gene is present in an intact form in the cell and can be expressed. Expression of one, some, or all of the HEPATOX sequences is then detected, if present, and, preferably, measured. Using sequence information provided by the database entries for the known sequences, or the sequence information for the newly described sequences, expression of the HEPATOX sequences can be detected (if present) and measured using techniques well known to one of ordinary skill in the art. For example, sequences within the sequence database entries corresponding to HEPATOX sequences, or within the sequences disclosed herein, can be used to construct probes for detecting HEPATOX RNA sequences in, *e.g.*, northern blot hybridization analyses or methods which specifically, and, preferably, quantitatively amplify specific nucleic acid sequences. As another example, the sequences can be used to construct primers for specifically amplifying

the HEPATOX sequences in, *e.g.*, amplification-based detection methods such as reverse-transcription based polymerase chain reaction. When alterations in gene expression are associated with gene amplification or deletion, sequence comparisons in test and reference populations can be made by comparing relative amounts of the examined DNA sequences in the test and reference cell populations.

Expression can be also measured at the protein level, *i.e.*, by measuring the levels of polypeptides encoded by the gene products described herein. Such methods are well known in the art and include, *e.g.*, immunoassays based on antibodies to proteins encoded by the genes.

Expression level of one or more of the HEPATOX sequences in the test cell population is then compared to expression levels of the sequences in one or more cells from a reference cell population. Expression of sequences in test and control populations of cells can be compared using any art-recognized method for comparing expression of nucleic acid sequences. For example, expression can be compared using GENECALLING® methods as described in US Patent No. 5,871,697 and in Shimkets et al., Nat. Biotechnol. 17:798-803.

In various embodiments, the expression of one or more sequences encoding genes of related function, as listed in Table 1, is compared. These functions include, *e.g.*, "Molecular Toxicology Genes" (such as, HEPATOX 22-24), "PPAR Ligands" (HEPATOX 25-34), "Ligand Activated Genes", "Acute Phase Protein Genes/Immune Response Genes" (HEPATOX 35-53), "Stress Response Genes" (HEPATOX 54-56), "Serum Protein Related Genes" (HEPATOX 57 and 58), and "Drug Metabolism and Free Radical Scavenger Genes" (HEPATOX 59-73). In some embodiments, expression of members of two or more functional families are compared.

In various embodiments, the expression of 2, 3, 4, 5, 6, 7,8, 9, 10, 15, 20, 25, 35, 40, 50, 100 or all of the sequences represented by HEPATOX 1-169 are measured. If desired, expression of these sequences can be measured along with other sequences whose expression is known to be altered according to one of the herein described parameters or conditions.

The reference cell population includes one or more cells for which the compared parameter is known. The compared parameter can be, *e.g.*, hepatotoxic agent expression status. By "hepatotoxic agent expression status" is meant that it is known whether the reference cell has had contact with a hepatotoxic agent. An example of a hepatotoxic agent is,

e.g., a thiazolidinedione such as troglitazone. Whether or not comparison of the gene expression profile in the test cell population to the reference cell population reveals the presence, or degree, of the measured parameter depends on the composition of the reference cell population. For example, if the reference cell population is composed of cells that have not been treated with a known hepatotoxic agent, a similar gene expression level in the test cell population and a reference cell population indicates the test agent is not a hepatotoxic agent. Conversely, if the reference cell population is made up of cells that have been treated with a hepatotoxic agent, a similar gene expression profile between the test cell population and the reference cell population indicates the test agent is a hepatotoxic agent.

In various embodiments, a HEPATOX sequence in a test cell population is considered comparable in expression level to the expression level of the HEPATOX sequence if its expression level varies within a factor of 2.0, 1.5, or 1.0 fold to the level of the HEPATOX transcript in the reference cell population. In various embodiments, a HEPATOX sequence in a test cell population can be considered altered in levels of expression if its expression level varies from the reference cell population by more than 1.0, 1.5, 2.0 or more fold from the expression level of the corresponding HEPATOX sequence in the reference cell population.

If desired, comparison of differentially expressed sequences between a test cell population and a reference cell population can be done with respect to a control nucleic acid whose expression is independent of the parameter or condition being measured. Expression levels of the control nucleic acid in the test and reference nucleic acid can be used to normalize signal levels in the compared populations.

In some embodiments, the test cell population is compared to multiple reference cell populations. Each of the multiple reference populations may differ in the known parameter. Thus, a test cell population may be compared to a first reference cell population known to have been exposed to a hepatotoxic agent, as well as a second reference population known have not been exposed to a hepatotoxic agent.

The test cell population that is exposed to, *i.e.*, contacted with, the test hepatotoxic agent can be any number of cells, *i.e.*, one or more cells, and can be provided *in vitro, in vivo*, or *ex vivo*.

In other embodiments, the test cell population can be divided into two or more subpopulations. The subpopulations can be created by dividing the first population of cells to create as identical a subpopulation as possible. This will be suitable, in, for example, *in vitro* or *ex vivo* screening methods. In some embodiments, various sub populations can be exposed to a control agent, and/or a test agent, multiple test agents, or, *e.g.*, varying dosages of one or multiple test agents administered together, or in various combinations.

Preferably, cells in the reference cell population are derived from a tissue type as similar as possible to test cell, *e.g.*, liver tissue. In some embodiments, the control cell is derived from the same subject as the test cell, *e.g.*, from a region proximal to the region of origin of the test cell. In other embodiments, the reference cell population is derived from a plurality of cells. For example, the reference cell population can be a database of expression patterns from previously tested cells for which one of the herein-described parameters or conditions (hepatotoxic agent expression status is known.

The test agent can be a compound not previously described or can be a previously known compound but which is not known to be a hepatotoxic agent.

The subject is preferably a mammal. The mammal can be, *e.g.*, a human, non-human primate, mouse, rat, dog, cat, horse, or cow.

SCREENING FOR TOXIC AGENTS

In one aspect, the invention provides a method of identifying toxic agents, *e.g.*, hepatotoxic agents. The hepatotoxic agent can be identified by providing a cell population that includes cells capable of expressing one or more nucleic acid sequences homologous to those listed in Table 1. as HEPATOX 1-169. The sequences need not be identical to sequences including HEPATOX 1-169, as long as the sequence is sufficiently similar that specific hybridization can be detected. Preferably, the cell includes sequences that are identical, or nearly identical to those identifying the HEPATOX nucleic acids shown in Table 1.

Expression of the nucleic acid sequences in the test cell population is then compared to the expression of the nucleic acid sequences in a reference cell population, which is a cell population that has not been exposed to the test agent, or, in some embodiments, a cell population exposed the test agent. Comparison can be performed on test and reference

samples measured concurrently or at temporally distinct times. An example of the latter is the use of compiled expression information, *e.g.*, a sequence database, which assembles information about expression levels of known sequences following administration of various agents. For example, alteration of expression levels following administration of test agent can be compared to the expression changes observed in the nucleic acid sequences following administration of a control agent, such as troglitazone.

An alteration in expression of the nucleic acid sequence in the test cell population compared to the expression of the nucleic acid sequence in the reference cell population that has not been exposed to the test agent indicates the test agent is a hepatotoxic agent.

The invention also includes a hepatotoxic agent identified according to this screening method.

ASSESSING TOXICITY OF AN AGENT IN A SUBJECT

The differentially expressed HEPATOX sequences identified herein also allow for the hepatotoxicity of a hepatotoxic agent to be determined or monitored. In this method, a test cell population from a subject is exposed to a test agent, *i.e.* a. hepatotoxic agent. If desired, test cell populations can be taken from the subject at various time points before, during, or after exposure to the test agent. Expression of one or more of the HEPATOX sequences, *e.g.*, HEPATOX: 1-169, in the cell population is then measured and compared to a reference cell population which includes cells whose hepatotoxic agent expression status is known. Preferably, the reference cells not been exposed to the test agent.

If the reference cell population contains no cells exposed to the treatment, a similarity in expression between HEPATOX sequences in the test cell population and the reference cell population indicates that the treatment is non-hepatotoxic. However, a difference in expression between HEPATOX sequences in the test population and this reference cell population indicates the treatment is hepatotoxic.

By "hepatotoxicity" is meant that the agent is damaging or destructive to liver when administered to a subject. leads to liver damage.

HEPATOX NUCLEIC ACIDS

Also provided in the invention are novel nucleic acid comprising a nucleic acid sequence selected from the group consisting of HEPATOX:1-21, or its complement, as well as vectors and cells including these nucleic acids.

Thus, one aspect of the invention pertains to isolated HEPATOX nucleic acid molecules that encode HEPATOX proteins or biologically active portions thereof. Also included are nucleic acid fragments sufficient for use as hybridization probes to identify HEPATOX-encoding nucleic acids (*e.g.*, HEPATOX mRNA) and fragments for use as polymerase chain reaction (PCR) primers for the amplification or mutation of HEPATOX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt) or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated HEPATOX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4

kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of any of HEPATOX:1-21, or a complement of any of these nucleotide sequences, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of these nucleic acid sequences as a hybridization probe, HEPATOX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to *HEPATOX* nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having at least about 10 nt and as many as 50 nt, preferably about 15 nt to 30 nt. They may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in HEPATOX:

1-21 . In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of these sequences, or a portion of any of these nucleotide sequences. A nucleic acid molecule that is complementary to the nucleotide sequence shown in HEPATOX:1-21 is one that is sufficiently complementary to the nucleotide sequence shown, such that it can hydrogen bond with little or no mismatches to the nucleotide sequences shown, thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of HEPATOX:1-21 *e.g.*, a fragment that can be used as a probe or primer or a fragment encoding a biologically active portion of HEPATOX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or

analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 45%, 50%, 70%, 80%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which in incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of a HEPATOX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a HEPATOX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding a human HEPATOX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in a HEPATOX polypeptide, as well as a polypeptide having a HEPATOX activity. A homologous amino acid sequence does not encode the amino acid sequence of a human HEPATOX polypeptide.

The nucleotide sequence determined from the cloning of human HEPATOX genes allows for the generation of probes and primers designed for use in identifying and/or cloning HEPATOX homologues in other cell types, *e.g.*, from other tissues, as well as HEPATOX homologues from other mammals. The probe/primer typically comprises a substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence of a nucleic acid comprising a HEPATOX sequence, or an anti-sense strand nucleotide sequence of a nucleic acid comprising a HEPATOX sequence, or of a naturally occurring mutant of these sequences.

Probes based on human HEPATOX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a HEPATOX protein, such as by measuring a level of a HEPATOX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting HEPATOX mRNA levels or determining whether a genomic HEPATOX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of HEPATOX" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of HEPATOX" can be prepared by isolating a portion of HEPATOX:1-21, that encodes a polypeptide having a HEPATOX biological activity, expressing the encoded portion of HEPATOX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of HEPATOX. For example, a nucleic acid fragment encoding a biologically active portion of a HEPATOX polypeptide can optionally include an ATP-binding domain. In another embodiment, a nucleic acid fragment encoding a biologically active portion of HEPATOX includes one or more regions.

HEPATOX VARIANTS

The invention further encompasses nucleic acid molecules that differ from the disclosed or referenced HEPATOX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same HEPATOX protein as that encoded by nucleotide sequence comprising a HEPATOX nucleic acid as shown in, *e.g.*, ADIPO1-21

In addition to the rat HEPATOX nucleotide sequence shown in HEPATOX:1-21, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of a HEPATOX polypeptide may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the HEPATOX gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a HEPATOX protein, preferably a mammalian HEPATOX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the HEPATOX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in HEPATOX that are the result of natural allelic variation and that do not alter the functional activity of HEPATOX are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding HEPATOX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of ADIPO1-21, are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the HEPATOX DNAs of the invention can be isolated based on their homology to the human HEPATOX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. For example, a soluble human HEPATOX DNA can be isolated based on its homology to human membrane-bound HEPATOX. Likewise, a membrane-bound human HEPATOX DNA can be isolated based on its homology to soluble human HEPATOX.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of HEPATOX:1-21. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250 or 500 nucleotides in length. In

another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding HEPATOX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at Tm, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g., 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic

acid molecule of the invention that hybridizes under stringent conditions to the sequence of HEPATOX:1-21 corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of HEPATOX:1-21 or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of HEPATOX:1-21or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.*, as employed for cross-species hybridizations). See, *e.g.*, Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo et al., 1981, *Proc Natl Acad Sci USA 78*: 6789-6792.

CONSERVATIVE MUTATIONS

In addition to naturally-occurring allelic variants of the HEPATOX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced into an HEPATOX nucleic acid or directly into an HEPATOX polypeptide sequence without altering the functional ability of the HEPATOX protein. In some

embodiments, the nucleotide sequence of HEPATOX:1-21will be altered, thereby leading to changes in the amino acid sequence of the encoded HEPATOX protein. For example, nucleotide substitutions that result in amino acid substitutions at various "non-essential" amino acid residues can be made in the sequence of HEPATOX:1-21. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of HEPATOX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the HEPATOX proteins of the present invention, are predicted to be particularly unamenable to alteration.

In addition, amino acid residues that are conserved among family members of the HEPATOX proteins of the present invention, are also predicted to be particularly unamenable to alteration. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the HEPATOX proteins) may not be essential for activity and thus are likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding HEPATOX proteins that contain changes in amino acid residues that are not essential for activity. Such HEPATOX proteins differ in amino acid sequence from the amino acid sequences of polypeptides encoded by nucleic acids containing HEPATOX:1-21, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous, more preferably 60%, and still more preferably at least about 70%, 80%, 90%, 95%, 98%, and most preferably at least about 99% homologous to the amino acid sequence of the amino acid sequences of polypeptides encoded by nucleic acids comprising HEPATOX:1-21.

An isolated nucleic acid molecule encoding a HEPATOX protein homologous to can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of a nucleic acid comprising HEPATOX:1-21, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into a nucleic acid comprising HEPATOX:1-21 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in HEPATOX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a HEPATOX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for HEPATOX biological activity to identify mutants that retain activity. Following mutagenesis of the nucleic acid, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant HEPATOX protein can be assayed for (1) the ability to form protein:protein interactions with other HEPATOX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant HEPATOX protein and a HEPATOX ligand; (3) the ability of a mutant HEPATOX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind ATP; or (5) the ability to specifically bind a HEPATOX protein antibody.

In other embodiment, the fragment of the complementary polynucleotide sequence described in claim 1 wherein the fragment of the complementary polynucleotide sequence hybridizes to the first sequence.

In other specific embodiments, the nucleic acid is RNA or DNA. The fragment or the fragment of the complementary polynucleotide sequence described in claim 38, wherein the fragment is between about 10 and about 100 nucleotides in length, *e.g.*, between about 10 and about 90 nucleotides in length, or about 10 and about 75 nucleotides in length, about 10 and

about 50 bases in length, about 10 and about 40 bases in length, or about 15 and about 30 bases in length.

ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of a HEPATOX sequence or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire HEPATOX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a HEPATOX protein, or antisense nucleic acids complementary to a nucleic acid comprising a HEPATOX nucleic acid sequence are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding HEPATOX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding HEPATOX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding HEPATOX disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of HEPATOX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of HEPATOX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of HEPATOX mRNA. An antisense oligonucleotide can be, for example,

about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a HEPATOX protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively,

antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave HEPATOX mRNA transcripts to thereby inhibit translation of HEPATOX mRNA. A ribozyme having specificity for a HEPATOX-encoding nucleic acid can be designed based upon the nucleotide sequence of a HEPATOX DNA disclosed herein. For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a HEPATOX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, HEPATOX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, HEPATOX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of a HEPATOX nucleic acid (*e.g.*, the HEPATOX promoter and/or enhancers) to form triple helical structures that prevent transcription of the HEPATOX gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of HEPATOX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of HEPATOX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of HEPATOX can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of HEPATOX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of HEPATOX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion

while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

HEPATOX POLYPEPTIDES

One aspect of the invention pertains to isolated HEPATOX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-HEPATOX antibodies. In one embodiment, native HEPATOX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, HEPATOX proteins are produced by recombinant DNA techniques. Alternative

to recombinant expression, a HEPATOX protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the HEPATOX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of HEPATOX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of HEPATOX protein having less than about 30% (by dry weight) of non-HEPATOX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-HEPATOX protein, still more preferably less than about 10% of non-HEPATOX protein, and most preferably less than about 5% non-HEPATOX protein. When the HEPATOX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of HEPATOX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of HEPATOX protein having less than about 30% (by dry weight) of chemical precursors or non-*HEPATOX* chemicals, more preferably less than about 20% chemical precursors or non-HEPATOX chemicals, still more preferably less than about 10% chemical precursors or non-HEPATOX chemicals, and most preferably less than about 5% chemical precursors or non-HEPATOX chemicals.

Biologically active portions of a HEPATOX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the HEPATOX protein, *e.g.*, the amino acid sequence encoded by a nucleic acid comprising HEPATOX 1-20 that include fewer amino acids than the full length HEPATOX proteins, and exhibit at least one activity of a HEPATOX protein. Typically, biologically active portions

comprise a domain or motif with at least one activity of the HEPATOX protein. A biologically active portion of a HEPATOX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a HEPATOX protein of the present invention may contain at least one of the above-identified domains conserved between the HEPATOX proteins. An alternative biologically active portion of a HEPATOX protein may contain at least two of the above-identified domains. Another biologically active portion of a HEPATOX protein may contain at least three of the above-identified domains. Yet another biologically active portion of a HEPATOX protein of the present invention may contain at least four of the above-identified domains.

Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native HEPATOX protein.

In some embodiments, the HEPATOX protein is substantially homologous to one of these HEPATOX proteins and retains its the functional activity, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below.

In specific embodiments, the invention includes an isolated polypeptide comprising an amino acid sequence that is 80% or more identical to the sequence of a polypeptide whose expression is modulated in a mammal to which hepatotoxic agent is administered.

DETERMINING HOMOLOGY BETWEEN TWO OR MORE SEQUENCES

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See *Needleman and Wunsch* 1970 *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of a DNA sequence comprising HEPATOX: 1-21.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison *(i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

CHIMERIC AND FUSION PROTEINS

The invention also provides HEPATOX chimeric or fusion proteins. As used herein, an HEPATOX "chimeric protein" or "fusion protein" comprises an HEPATOX polypeptide operatively linked to a non-HEPATOX polypeptide. A "HEPATOX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to HEPATOX, whereas a "non-HEPATOX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the HEPATOX protein, *e.g.*, a protein that is different from the HEPATOX protein and that is derived from the same or a different organism. Within an HEPATOX fusion protein the HEPATOX polypeptide can

correspond to all or a portion of an HEPATOX protein. In one embodiment, an HEPATOX fusion protein comprises at least one biologically active portion of an HEPATOX protein. In another embodiment, an HEPATOX fusion protein comprises at least two biologically active portions of an HEPATOX protein. In yet another embodiment, an HEPATOX fusion protein comprises at least three biologically active portions of an HEPATOX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the HEPATOX polypeptide and the non-HEPATOX polypeptide are fused in-frame to each other. The non-HEPATOX polypeptide can be fused to the N-terminus or C-terminus of the HEPATOX polypeptide.

For example, in one embodiment an HEPATOX fusion protein comprises an HEPATOX domain operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds which modulate HEPATOX activity (such assays are described in detail below).

In yet another embodiment, the fusion protein is a GST-HEPATOX fusion protein in which the HEPATOX sequences are fused to the C-terminus of the GST (*i.e.*, glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant HEPATOX.

In another embodiment, the fusion protein is an HEPATOX protein containing a heterologous signal sequence at its N-terminus. For example, a native HEPATOX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of HEPATOX can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is an HEPATOX-immunoglobulin fusion protein in which the HEPATOX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The HEPATOX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a HEPATOX ligand and a HEPATOX protein on the surface of a cell, to thereby suppress HEPATOX-mediated signal transduction *in vivo*. The HEPATOX-immunoglobulin fusion proteins can be used to affect the bioavailability of an HEPATOX cognate ligand. Inhibition

of the HEPATOX ligand/HEPATOX interaction may be useful therapeutically for both the treatments of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the HEPATOX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-HEPATOX antibodies in a subject, to purify HEPATOX ligands, and in screening assays to identify molecules that inhibit the interaction of HEPATOX with a HEPATOX ligand.

An HEPATOX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel *et al.* (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An HEPATOX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the HEPATOX protein.

HEPATOX AGONISTS AND ANTAGONISTS

The present invention also pertains to variants of the HEPATOX proteins that function as either HEPATOX agonists (mimetics) or as HEPATOX antagonists. Variants of the HEPATOX protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the HEPATOX protein. An agonist of the HEPATOX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the HEPATOX protein. An antagonist of the HEPATOX protein can inhibit one or more of the activities of the naturally occurring form of the HEPATOX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade

which includes the HEPATOX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the HEPATOX proteins.

Variants of the HEPATOX protein that function as either HEPATOX agonists (mimetics) or as HEPATOX antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the HEPATOX protein for HEPATOX protein agonist or antagonist activity. In one embodiment, a variegated library of HEPATOX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of HEPATOX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential HEPATOX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of HEPATOX sequences therein. There are a variety of methods which can be used to produce libraries of potential HEPATOX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential HEPATOX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477.

POLYPEPTIDE LIBRARIES

In addition, libraries of fragments of the HEPATOX protein coding sequence can be used to generate a variegated population of HEPATOX fragments for screening and subsequent selection of variants of an HEPATOX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a HEPATOX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form

double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the HEPATOX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of HEPATOX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify HEPATOX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

ANTI-HEPATOX ANTIBODIES

An isolated HEPATOX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind HEPATOX using standard techniques for polyclonal and monoclonal antibody preparation. The full-length HEPATOX protein can be used or, alternatively, the invention provides antigenic peptide fragments of HEPATOX for use as immunogens. The antigenic peptide of HEPATOX comprises at least 8 amino acid residues of the amino acid sequence encoded by a nucleic acid comprising the nucleic acid sequence shown in HEPATOX:1-21 and encompasses an epitope of HEPATOX such that an antibody raised against the peptide forms a specific immune complex with HEPATOX. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of HEPATOX that are located on the surface of the protein, *e.g.*,

hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

HEPATOX polypeptides or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab')2}$ fragments, and an F_{ab} expression library. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to an HEPATOX protein sequence, or derivatives, fragments, analogs or homologs thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (*e.g.*, rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed HEPATOX protein or a chemically synthesized HEPATOX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (*e.g.*, aluminum hydroxide), surface active substances (*e.g.*, lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against HEPATOX can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen

binding site capable of immunoreacting with a particular epitope of HEPATOX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular HEPATOX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular HEPATOX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a HEPATOX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*, 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a HEPATOX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Antibody fragments that contain the idiotypes to a HEPATOX protein may be produced by techniques known in the art including, but not limited to: (*i*) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (*ii*) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (*iii*) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (*iv*) F_v fragments.

Additionally, recombinant anti-HEPATOX antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT

International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.*(1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988) *J Natl Cancer Inst.* 80:1553-1559); Morrison(1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a HEPATOX protein is facilitated by generation of hybridomas that bind to the fragment of a HEPATOX protein possessing such a domain. Antibodies that are specific for one or more domains within a HEPATOX protein, *e.g.,* domains spanning the above-identified conserved regions of HEPATOX family proteins, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-HEPATOX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a HEPATOX protein (*e.g.,* for use in measuring levels of the HEPATOX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for HEPATOX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-HEPATOX antibody (*e.g.*, monoclonal antibody) can be used to isolate HEPATOX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-HEPATOX antibody can facilitate the purification of natural HEPATOX from cells and of recombinantly produced HEPATOX expressed in host cells. Moreover, an anti-HEPATOX antibody can be used to detect HEPATOX protein (*e.g.*, in a cellular lysate or

cell supernatant) in order to evaluate the abundance and pattern of expression of the HEPATOX protein. Anti-HEPATOX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

HEPATOX RECOMBINANT EXPRESSION VECTORS AND HOST CELLS

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding HEPATOX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a linear or circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such

other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, HEPATOX proteins, mutant forms of HEPATOX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of HEPATOX in prokaryotic or eukaryotic cells. For example, HEPATOX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or

non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the HEPATOX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, HEPATOX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That

is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to HEPATOX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, HEPATOX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory,

Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding HEPATOX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) an HEPATOX protein. Accordingly, the invention further provides methods for producing HEPATOX protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding HEPATOX has been introduced) in a suitable medium such that HEPATOX protein is produced. In another embodiment, the method further comprises isolating HEPATOX from the medium or the host cell.

PHARMACEUTICAL COMPOSITIONS

The HEPATOX nucleic acid molecules, HEPATOX proteins, and anti-HEPATOX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or

diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and

antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a HEPATOX protein or anti-HEPATOX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

KITS AND NUCLEIC ACID COLLECTIONS FOR IDENTIFYING HEPATOX NUCLEIC ACIDS

In another aspect, the invention provides a kit useful for examining hepatotoxicity of agents. The kit can include nucleic acids that detect two or more HEPATOX sequences. In preferred embodiments, the kit includes reagents which detect 3, 4, 5, 6, 8, 10, 12, 15, 20, 25, 50, 100 or all of the HEPATOX nucleic acid sequences.

The invention also includes an isolated plurality of sequences which can identify one or more HEPATOX responsive nucleic acid sequences.

The kit or plurality may include, *e.g.*, sequence homologous to HEPATOX nucleic acid sequences, or sequences which can specifically identify one or more HEPATOX nucleic acid sequences.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

What is claimed is:

1. A method of screening a test agent for hepatotoxicity, the method comprising;

(a) providing a test cell population comprising a cell capable of expressing one or more nucleic acid sequences selected from the group consisting of HEPATOX: 1-168 and 169;

(b) contacting the test cell population with a test agent;

(c) measuring expression of one or more of the nucleic acid sequences in the test cell population;

(d) comparing the expression of the nucleic acid sequences in the test cell population to the expression of the nucleic acid sequences in a reference cell population comprising at least one cell whose exposure status to a hepatotoxic agent is known; and

(e) identifying a difference in expression levels of the HEPATOX sequence, if present, in the test cell population and reference cell population,

thereby screening said test agent for hepatotoxicity.

2. The method of claim 1, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of HEPATOX 22- 24 and HEPATOX59-73.

3. The method of claim 1, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of HEPATOX 25-34.

4. The method of claim 2, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of HEPATOX 22-24 and HEPATOX59-73.

5. The method of claim 1, wherein the method comprises comparing the expression of 40 or more of the nucleic acid sequences.

6. The method of claim 1, wherein the expression of the nucleic acid sequences in the test cell population is decreased as compared to the reference cell population.

7. The method of claim 1, wherein the expression of the nucleic acid sequences in the test cell population is increased as compared to the reference cell population.

8. The method of claim 1, wherein the test cell population is provided *in vitro*.

9. The method of claim 1, wherein the test cell population is provided *ex vivo* from a mammalian subject.

10. The method of claim 1, wherein the test cell population is provided *in vivo* in a mammalian subject.

11. The method of claim 1, wherein the test cell population is derived from a human or rodent subject.

12. The method of claim 1, wherein the test cell population includes a hepatocyte.

13. The method of claim 1, wherein said test agent is a ligand for a PPARγ receptor.

14. The method of claim 1, wherein the hepatotoxic agent is a thiazolidinedione.

15. The method of claim 14, wherein the thiazolidinedione is troglitazone.

16. A method of assessing the hepatotoxicity of a test agent in a subject, the method comprising:

(a) providing from the subject a test cell population comprising a cell capable of expressing one or more nucleic acid sequences selected from the group consisting of HEPATOX: 1-168 and 169;

(b) contacting the test cell population with a test agent;

(c) measuring expression of one or more of the nucleic acid sequences in the test cell population; and

(d) comparing the expression of the nucleic acid sequences in the test cell population to the expression of the nucleic acid sequences in a reference cell population comprising at least one cell whose exposure status to a hepatotoxic agent is known;

(e) identifying a difference in expression levels of the nucleic acid sequences, if present, in the test cell population and the reference cell population,

thereby assessing the hepatotoxicity of the test agent in the subject.

17. The method of claim 16, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of and HEPATOX59-73.

18. The method of claim 16, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of HEPATOX 25-34.

19. The method of claim 18, wherein the method comprises comparing the expression of one or more genes selected from the group consisting of and HEPATOX59-73.

20. The method of claim 16, wherein the expression of the nucleic acid sequences in the test cell population is decreased as compared to the reference cell population.

21. The method of claim 16, wherein the expression of the nucleic acid sequences in the test cell population is increased as compared to the reference cell population.

22. The method of claim 16, wherein said subject is a human or rodent.

23. The method of claim 16, wherein the test cell population is provided *ex vivo* from said subject.

24. The method of claim 16, wherein the test cell population is provided *in vivo* from said subject.

25. An isolated nucleic acid comprising a nucleic acid sequence selected from the group consisting of a HEPATOX 1-21 nucleic acid, or its complement.

26. A vector comprising the nucleic acid of claim 25.

27. A cell comprising the vector of claim 26.

28. A pharmaceutical composition comprising the nucleic acid of claim 25.

29. A polypeptide encoded by the nucleic acid of claim 25.

30. An antibody which specifically binds to the polypeptide of claim 29.

31. A kit which detects two or more of the nucleic acid sequences selected from the group consisting of HEPATOX: 1-168 and 169.

32. An array which detects one or more of the nucleic acid selected from the group consisting of HEPATOX: 1-168 and 169.

33. A plurality of nucleic acid comprising one or more of the nucleic acid selected from the group consisting of HEPATOX: 1-168 and 169.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 October 2000 (26.10.2000)

PCT

(10) International Publication Number
WO 00/063435 A3

(51) International Patent Classification[7]: C12Q 1/68, C12N 15/70, 5/10, C07K 14/47, 16/18

(21) International Application Number: PCT/US00/10076

(22) International Filing Date: 14 April 2000 (14.04.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/129,763	15 April 1999 (15.04.1999)	US
60/156,924	28 September 1999 (28.09.1999)	US
09/548,589	13 April 2000 (13.04.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/129,763 (CIP)
Filed on	15 April 1999 (15.04.1999)
US	60/156,924 (CIP)
Filed on	28 September 1999 (28.09.1999)
US	09/548,589 (CIP)
Filed on	13 April 2000 (13.04.2000)

(71) Applicant *(for all designated States except US)*: **CURAGEN CORPORATION** [US/US]; 11th Floor, 555 Long Wharf Drive, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GOULD-ROTHBERG, Bonnie, E.** [CA/US]; 1701 Moose Hill Road, Guilford, CT 06437 (US). **DIPIPPO, Vincent, A.** [US/US]; 156 North Atwater Street, East Haven, CT 06512 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovksy, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
12 September 2002

(15) Information about Correction:
Previous Correction:
see PCT Gazette No. 19/2001 of 10 May 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/063435 A3

(54) Title: METHOD OF IDENTIFYING TOXIC AGENTS USING DIFFERENTIAL GENE EXPRESSION

(57) Abstract: Disclosed are methods of identifying toxic agents, *e.g.*, hepatotoxic agents, using differential gene expression. Also disclosed are novel nucleic acid sequences whose expression is differentially regulated by troglitazone.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/10076

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C12N15/70 C12N5/10 C07K14/47 C07K16/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS, CHEM ABS Data, MEDLINE, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 97 13877 A (LYNX THERAPEUTICS INC ;MARTIN DAVID W (US)) 17 April 1997 (1997-04-17) page 7, line 18 -page 43; claims 1-30 ---	1,5-16, 20-33
X	WO 98 13353 A (AURORA BIOSCIENCES CORP ;WHITNEY MICHAEL A (US); CRAIG FRANK (US);) 2 April 1998 (1998-04-02) page 34, line 17 -page 40, line 10 page 26, line 24 --- -/--	1,5-16, 20-33

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 October 2001	16. 01. 2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gabriels, J

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 00/10076

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE BIOSIS [Online] BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 1 July 1998 (1998-07-01) NEUSCHWANDER-TETRI BRENT A ET AL: "Troglitazone-induced hepatic failure leading to liver transplantation. A case report." Database accession no. PREV199800354952 XP002179971 abstract & ANNALS OF INTERNAL MEDICINE, vol. 129, no. 1, 1 July 1998 (1998-07-01), pages 38-41, ISSN: 0003-4819	14,15
E	WO 01 02609 A (DIPIPPO VINCENT A ;GOULD ROTHBERG BONNIE E (US); CURAGEN CORP (US)) 11 January 2001 (2001-01-11) the whole document	1,5-16, 20-33
E	WO 01 53514 A (GLAXO GROUP LTD ;REIDHAAR OLSON JOHN F (US)) 26 July 2001 (2001-07-26) claims 1-28	1,5-16, 20-33
E	WO 01 51638 A (INCYTE GENOMICS INC ;AZIMZAI YALDA (US); REDDY ROOPA (US); RING HU) 19 July 2001 (2001-07-19) claim 28	1,5-16, 20-33
E	WO 01 32928 A (FARR SPENCER ;PHASE 1 MOLECULAR TOXICOLOGY (US)) 10 May 2001 (2001-05-10) claims 1-39	1,5-16, 20-33

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/10076

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1, 5-16, 20-33 (all partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: Invention 1: claims 1,5-16,20-33 (all partially)

   An isolated nucleic acid comprising HEPATOX:1 or its complement, a polypeptide encoded thereby, specific antibodies thereto, vectors and cells derived thereof for a method of screening a test agent for hepatoxicity. Kits and arrays for detecting the expression profile of HEPATOX:1.

2. Claims: Invention 2-21: claims 1,5-16,20-33 (all partially)

   Invention 2:

   An isolated nucleic acid comprising HEPATOX:2 or its complement, a polypeptide encoded thereby, specific antibodies thereto, vectors and cells derived thereof for a method of screening a test agent for hepatoxicity. Kits and arrays for detecting the expression profile of HEPATOX:2.

   Invention 3-21:

   Idem for invention 3-21 but limited to HEPATOX:3-21.

3. Claims: Invention 22: claims 1-24,31-33 (all partially)

   A method for screening a test agent for hepatoxicity using the expression profile of HEPATOX:22. Kits and arrays for detecting the expression profile of HEPATOX:22.

4. Claims: Invention 23-169: claims 1-24,31-33 (all partially)

   Invention 23:

   A method for screening a test agent for hepatoxicity using the expression profile of HEPATOX:23. Kits and arrays for detecting the expression profile of HEPATOX:23.

   Invention 24-169:

   Idem for inventions 24-169 but limited to HEPATOX:24-169.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 00/10076

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9713877 A	17-04-1997	AU 712929 B2	18-11-1999
		AU 4277896 A	06-05-1996
		AU 6102096 A	30-12-1996
		AU 7717596 A	30-04-1997
		CN 1193357 A	16-09-1998
		CZ 9700866 A3	17-09-1997
		CZ 9703926 A3	17-06-1998
		EP 0793718 A1	10-09-1997
		EP 0832287 A1	01-04-1998
		EP 0931165 A1	28-07-1999
		FI 971473 A	04-06-1997
		HU 9900910 A2	28-07-1999
		JP 11507528 T	06-07-1999
		JP 10507357 T	21-07-1998
		NO 971644 A	02-06-1997
		NO 975744 A	05-02-1998
		PL 324000 A1	27-04-1998
		US 6280935 B1	28-08-2001
		WO 9641011 A1	19-12-1996
		WO 9713877 A1	17-04-1997
		US 6235475 B1	22-05-2001
		AU 718357 B2	13-04-2000
WO 9813353 A	02-04-1998	US 5928888 A	27-07-1999
		AU 4505797 A	17-04-1998
		EP 0952976 A1	03-11-1999
		WO 9813353 A1	02-04-1998
		US 6107477 A	22-08-2000
WO 0102609 A	11-01-2001	AU 7133000 A	22-01-2001
		WO 0102609 A2	11-01-2001
WO 0153514 A	26-07-2001	AU 2796701 A	31-07-2001
		EP 1165825 A1	02-01-2002
		WO 0153514 A1	26-07-2001
WO 0151638 A	19-07-2001	AU 2790001 A	24-07-2001
		WO 0151638 A2	19-07-2001
WO 0132928 A	10-05-2001	AU 1466001 A	14-05-2001
		WO 0132928 A2	10-05-2001